Automatic Detection of Microlensing Events in the Galactic Bulge using Machine Learning Techniques

Selina Chu,[1], Kiri L. Wagstaff[1], Geoffrey Bryden[1] and Yossi Shvartzvald[2]

[1]*JPL, Caltech, Pasadena, CA, USA;* Selina.Chu@jpl.nasa.gov

[2]*IPAC, California Institute of Technology, Pasadena, CA, USA*

Abstract. The Wide Field Infrared Survey Telescope (WFIRST) is a NASA flagship mission scheduled to launch in mid-2020, with more than one year of its lifetime dedicated to microlensing survey. The survey is to discover thousands of exoplanets near or beyond the snowline via their microlensing light curve signatures, enabling a Kepler-like statistical analysis of planets at ~1-10 AU from their host stars and potentially revolutionizing our understanding of planet formation. The goal of our work is to create an automated system that has the ability to efficiently process and classify large-scale astronomical datasets that missions such as WFIRST will produce. In this paper, we discuss our framework that utilizes feature selection and parameter optimization for classification models to automatically differentiate the different types of stellar variability and detect microlensing events.

1. Introduction

Microlensing is an important technique for exoplanet detection and characterization. For a microlensing survey to be successful, the number of detected planets is proportional to the number of microlensing events, which in turn depends on the density of observable stars. The Galactic bulge is where the stellar surface density is highest. So naturally, microlensing surveys should concentrate their efforts toward the bulge to maximize the event rate. Traditionally, these surveys have been conducted at optical wavelengths, which suffer from high dust extinction near the Galactic bulge. Observing in the near-infrared (NIR) will mitigates the effects of high extinction, enabling observations closer to the Galactic center. Therefore to understand this potential, WFIRST will conduct its microlensing survey in the NIR (1-2μm) . However, until recently there has been little or no effort dedicated to microlensing surveys in the NIR. This means we do not have mapping of the microlensing event rate near the galactic center in the NIR, which makes it impossible to properly optimize WFIRST's science yield. For this reason, NASA's Exoplanet Program Analyis Group (ExoPAG) identified a NIR microlensing survey as a key precursor activity for WFIRST Yee et al. (2014). The goal of our work is to directly address these issues by determining the optimal target fields for the WFIRST microlensing survey and developing data analysis tools to enhance the science return of the survey. We propose a framework that could efficiently process lightcurves extracted from the tens of millions of stars in a NIR survey and fully automate identification of microlensing events. In this paper, we will describe our approach in developing a predictive model using machine learning to detect microlensing events

and characterize properties of these events. The use of feature selection enables us to learn which characteristics distinguish the different types of events and to classify high-dimensional data more efficiently. We demonstrate our proposed method on datasets acquired from UKIRT's wide-field near-IR camera that surveys the galactic bulge.

2. Microlensing Survey

To study the detection of microlensing events, we started a NIR survey with the United Kingdom Infrared Telescope (UKIRT), a 3.8-m telescope on Mauna Kea in Hawaii for our investigation. UKIRT was initially started as a pilot study in support of the 2015 Spitzer microlensing campaign, and in 2017, the program was redirected to cover all potential WFIRST fields, including the Galactic center. The full catalog of lightcurve data collected during the UKIRT microlensing campaigns is publicly available in NASA Exoplanet Archive.[1] Shvartzvald et al. (2017) has shown to have successfully identified the first five microlensing events in the NIR based on preliminary analysis on the 2016 UKIRT data.

To represent each lightcurve, we derived features using a grid-based approach for microlensing fit, based on a method proposed in Kim et al. (2017). The model grid utilizes the effective event timescale t_{eff} and the event peak time t_0, and each model lightcurve is scaled by the source flux F_s and blended flux F_b to derive analytically the best fit for each grid model. We utilize Markov Chain Monte Carlo (MCMC) techniques to fit (approximately) large number of llightcurves more efficiently Foreman-Mackey et al. (2013). is used to approximate fits for large number of llightcurves more efficiently. We initially filter the survey data to identify lightcurves that exhibit evidence for microlensing by comparing each microlensing fit against a straight line, providing their goodness of fit $\Delta_{\chi^2} \equiv (\chi^2_{flat} - \chi^2_{microlens})/(\chi^2_{microlens}/N_{points} - 4)$, where $\chi^2_{microlens}$ is the microlensing fit and N_{points} is the number of points in a lightcurve. Δ_{χ^2} furnishes us with a simple way to select lightcurves with significant variability as potential microlensing candidates. Table 1 describes the list of features derived from evaluating each lightcurve.

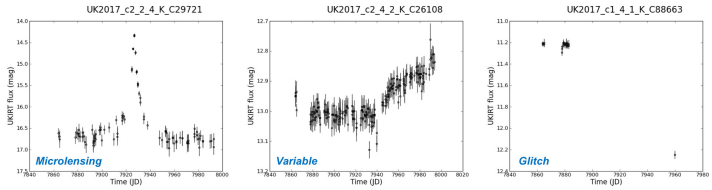

Figure 1. Examples of lightcurves *Left:* Microlensing, *Center:* Variable, *Left:* Glitch

3. Approach

The first step in building models for prediction is to obtain a set of groundtruths to train our classifier on. To accomplish this task, we manually label each observation by visual inspection using a The survey data was initially filtered using $\Delta_{\chi^2} > 100$, which resulted in a subset of approximately 30,000 potential microlensing candidates. From this subset, we manually labeled about 1,587 lightcurves into three types of events (or classes):

[1] `https://exoplanetarchive.ipac.caltech.edu/docs/UKIRTMission.html`

microlensing, intrinsic stellar variability events*microlensing*), and spurious instrumental artifacts (*glitches*). This results in 137 microlensing events, 1,083 variables, and 367 glitches. The rest of the unlabeled unlabeled lightcurves are used as test candidates for detection. Fig 1 depicts examples of the lightcurve from each of these events.

3.1. Classification with Feature Selection and Model Selection

We investigate on building a classifier using features listed in Table 1 with a focus on identifying microlensing events more effectively. To evaluate the performance of various classification algorithms to include in our recognition system, we examined three classification methods: *Random Forest* (RF), and *Support Vector Machine* (SVM), *K-Nearest Neighbor* (kNN), along with feature selection and model selection to optimize the classifiers' parameters.

When using a large number of features, there might be potentially irrelevant features that could negatively impact the quality of classification. Adding more features is not always helpful; as the number of features increases, the number of dimensions in the search space also increases, resulting in the data points becoming more sparse. We use feature selection to choose a more effective subset of features, which can reduce the overall computational cost and running time, as well as achieve an acceptable, if not higher, recognition rate. Instead of performing an exhaustive search of all features, we use a greedy backward elimination feature selection algorithm for our experiments. To perform model selection, we used grid search and varied the model parameters for each classifier, such as varying the number of trees in a forest, depth of each tree, kernel function, regularization constants, model constraints, etc. Model selection is incorporated at each iteration of the feature selection process to learn the optimal parameters for each classifier type and feature set. Our approach is developed based on the classifiers and model selection methods from Scikit-learn and feature selection from MLxtend.[2]

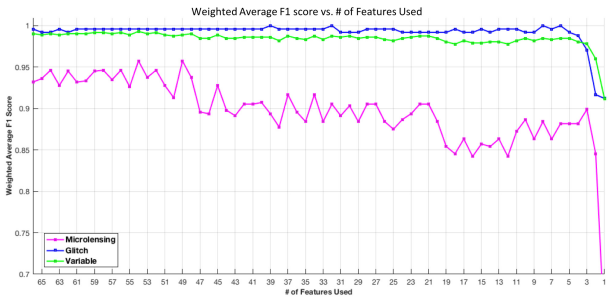

Figure 2. The effect classification F1-score as the number of features are removed.

For our experiment, we used the filtered dataset from 2017 UKIRT survey, and trained the classifiers to differentiate between*microlensing*, *variable*, and *glitch*. We used 3-fold cross-validation for model selection and feature selection. Data are normalized using zero mean and unit variance. F1-score was used for evaluation metric. F1-score is the harmonic average of the precision and recall, defined as $F1 = 2/(recall^{-1} + precision^{-1})$. Fig 2 shows the plots of the F1-score with respect to each class, as the number of features decreases. To obtain the F1-score for each class, the

[2]Scikit-learn: `scikit-learn.org`, MLxtend: `http://rasbt.github.io/mlxtend`

averaged F1-score is weighted relative to the amount of actual observations in each class. The classification results are summarized in Fig.3. We found results using SVM and RF to be similar with RF performing slightly better overall. To demonstrate the performance of our detector, we focus on using RF. The selected model parameters Θ_s utilizes 500 number of trees in a forest, max. depth of 4 for each tree, entropy criteria, and max $log_2(N_{features})$ features at each split. Features listed in Table 1 with (**) were pruned in the feature selection process. Fig 4 illustrates the lightcurves detected as microlensing event , from the set of unlabeled test candidates, using Θ_s with class probability (of being microlensing) $p_m(x) \geq 0.8$. Each lightcurve plot is ranked from highest to lowest (left to right, top to bottom).

(a) *Overall* 3-class classification

perc (%)	kNN	SVM	RF
Full Features	97.3	98.8	98.8
Selected Features	97.7	98.9	99.2

(b) *Microlensing* only

perc (%)	kNN	SVM	RF
Full Features	88.4	95.3	95.6
Selected Features	93.2	95.6	96.3

(c) Precision & Recall for *Microlensing* only

perc (%)	Full Features	Selected Features
Precision	89.8	93.8
Recall	95.6	97.8
False Positive	0.94	0.57
False Negative	0.38	0.19

Figure 3. Examples of detected microlensing events from unlabeled test candidates, ranked by highest to lowest $p_m(x)$ from left to right, top to bottom.

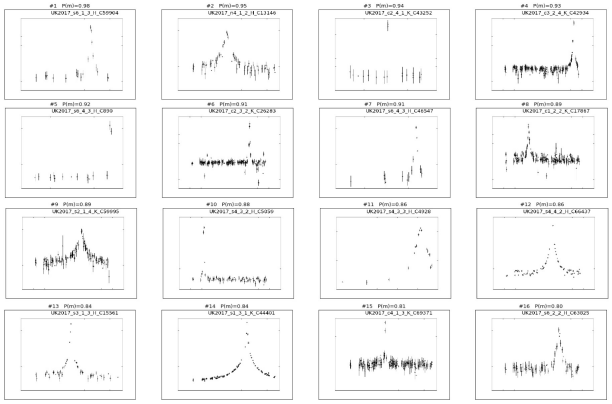

Figure 4. Examples of detected microlensing events from unlabeled test candidates, ranked by highest to lowest $p_m(x)$ from left to right, top to bottom.

4. Conclusions and Future Work

This paper investigates techniques for developing a microlensing detection pipeline using the proposed features. The classification system was successful in classifying the different types of events. We also found that using high number of features is not always beneficial to classification. In using forward feature selection we were able to achieve a slightly higher recognition rate, improving the overall classification. Currently we are developing a framework for injecting mock stars with microlensing signals into the UKIRT images to evaluate detection efficiency. Our ongoing work includes incorporating active learning into our system to improve classification performance of lightcurves by automatically selecting most informative unlabeled lightcurves, visually label the selected lightcurve, and then re-train our prediction models. To assure our ability to transfer our findings from UKIRT to WFIRST effectively, we plan to incorporate domain adaptation to bridge any differences between the two surveys.

Table 1. List of features. (**) indicates exclusion from *Selected features* set for exp. in Fig 4

Feature	Description
A_{sin}	amplitude of sinusoidal fit
σ_A	fractional uncertainty in A
b	intercept of linear fit
σ_b	uncertainty in b
$\chi^2_{baseline}$	χ^2 for a constant-flux fit
$\chi^2_{microlens}$ (**)	χ^2 for microlensing (mcmc fit)
χ^2_{drop1}	χ^2 for a constant-flux fit, dropping the worst point
χ^2_{drop2} (**)	χ^2 for a constant-flux fit, dropping two worst points
$\chi^2_{microlens}$	χ^2 for microlensing (mcmc fit)
$\chi^2_{reduced}$	χ^2 for microlensing fit / degrees of freedom
χ^2_{sin}	χ^2 for a sinusoidal fit
χ^2_{linear}	χ^2 for a linear fit
$\Delta\chi^2_{grid}$	χ^2 for microlensing (grid fit)
$\Delta\chi^2_{drop2}$	significance level for drop-2-points vs constant-flux
$\Delta\chi^2_{drop1,drop2}$	significance level for drop-1-point vs drop-2-points
$\Delta\chi^2_{microlens}$	significance level for microlensing (mcmc fit) vs constant-flux fit
$\Delta\chi^2_{sin,ulens}$	significance level for sinusoidal vs microlensing fit
$\Delta\chi^2_{sin}$ (**)	significance level for sinusoidal vs constant-flux fit
$\Delta\chi^2_{linear,ulens}$	significance level for linear vs microlensing fit
$\Delta\chi^2_{linear}$	significance level for linear vs baseline fit
σ	flux dispersion
ΔF_{data} (**)	range of fluxes within the data
ΔF_{model} (**)	rrange of fluxes within the model fit
F_{sin}	median flux of sinusoidal fit
σ_F (**)	uncertainty in F_{sin}
F_{max}	maximum observed flux
F_{tot} (**)	source flux (mcmc fit)
σ_{Ftot}	uncertainty in F_{tot}
f_1	first flux parameter (grid fit)
f_2	second flux parameter (grid fit)

Feature	Description
f_b	blending fraction (mcmc fit)
σ_{fb}	uncertainty in f_b
m	slope of linear fit
σ_m	uncertainty in m
$model$	0 or 1, depending on grid-fit model regime (small/large u_0)
N (**)	number of successful observations
N_{frac}	fraction of successful observations
N_{max}	number of observational epochs
$N_{high,2-\sigma}$ (**)	number of points 2-σ above the baseline
$f_{high,2-\sigma}$ (**)	fraction of points 2-σ above the baseline
$N_{high,3-\sigma}$	number of points 3-σ above the baseline
$f_{high,3-\sigma}$ (**)	fraction of points 3-σ above the baseline
P_{sin}	period of sinusoidal fit
σ_P	uncertainty in P_{sin}
σ_P/P	fractional uncertainty for period of sinusoidal fit
P/σ_P	S/N for period of sinusoidal fit
T_{sin} (**)	phase of sinusoidal fit
σ_T	uncertainty in T_{sin}
t_E	microlensing event timescale (mcmc fit)
$\lvert t_0 - t_{mid}\rvert - \Delta t/2$ (**)	how far event falls from edge of observing window
$(\lvert t_0 - t_{mid}\rvert - \Delta t/2)/t_E$	how far event falls from edge of observing window
σ_{tE}/t_E	fractional uncertainty in t_E
σ_{tE} (**)	uncertainty in t_E
t_E/σ_{tE} (**)	S/N for t_E
t_0 (**)	time of the microlensing event (mcmc fit)
σ_{to}	uncertainty in t_0
$t_{0,grid}$	time of the event (grid fit)
$t_0 - t_{mid}$	how well event falls within observing window
$(t_0 - t_{mid})/t_E$	how well event falls within observing window
$\lvert t_0 - t_{mid}\rvert$	how well event falls within observing window
$\lvert t_0 - t_{mid}\rvert/t_E$ (**)	how well event falls within observing window
t_{eff}	microlensing event timescale (mcmc fit)
u_0	microlensing event impact parameter (mcmc fit)
σ_{uo}/u_0	fractional uncertainty in u_0
σ_{uo}	uncertainty in u_0
u_0/σ_{uo}	S/N for u_0

5. References

References

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306. 1202.3665

Kim, D. J., Kim, H. W., Hwang, K. H., Albrow, M. D., Chung, S. J., Gould, A., Han, C., Jung, Y. K., Ryu, Y. H., Shin, I. G., Yee, J. C., Zhu, W., Cha, S. M., Kim, S. L., Lee, C. U., Lee, Y., Park, B. G., & Pogge, R. W. 2017, 1. 1703.06883

Shvartzvald, Y., Bryden, G., Gould, A., Henderson, C. B., Howell, S. B., & Beichman, C. 2017, AJ, 153, 61. 1610.02039

Yee, J. C., Albrow, M., Barry, R. K., Bryden, G., Chung, S.-j., Gaudi, B. S., Gehrels, N., Gould, A., Penny, M. T., Rattenbury, N., Skowron, J., Street, R., & Sumi, T. 2014, 1

Acknowledgments. The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.